AD-A051 394 TRACOR INC SAN DIEGO CALIF F/G 5/9
TOWARD THE DEVELOPMENT OF A SYSTEMATIC FUNCTIONAL METHODOLOGY F--ETC(U)
JUL 67 F J BLAISDELL
UNCLASSIFIED TRACOR-SD-67-016-U NL

| OF |
AD
A051394

END
DATE
FILMED
4 -78
DDC

Document Number
TRACOR SD-67-016-U

9

TECHNICAL MEMORANDUM

TOWARD THE DEVELOPMENT OF A SYSTEMATIC FUNCTIONAL METHODOLOGY FOR COMPUTER-ASSISTED INSTRUCTION

July 1967

SCIENCES AND SYSTEMS DIVISION • SAN DIEGO LABORATORY

TRACOR™

3065 Rosecrans Place, San Diego, California 92110, AC 714/222-6692
Home Office: TRACOR Inc., 6400 Tracor Lane, Austin, Texas 78721, AC 512-926-2800

TRACOR, INC. 3065 ROSECRANS PLACE, SAN DIEGO, CALIFORNIA 92110

14 Document Number
TRACOR-SD-67-016-U

9 TECHNICAL MEMORANDUM

6 Toward the Development of a Systematic
Functional Methodology for
Computer-Assisted Instruction,

11 July 1967

12 64p.

Approved:

Bruce H Deatherage
B. H. Deatherage, Ph. D.
Manager, Psychology
Section

Prepared by:

F. J. Blaisdell
10 F. J. Blaisdell Ph. D.
Group Leader, Personnel
and Training Laboratory,
Psychology Section

403 906 4B

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

LIST OF FIGURES

TRACOR, INC. 3065 ROSECRANS PLACE, SAN DIEGO, CALIFORNIA 92110

ABSTRACT

The memorandum presents a systematic functional methodology for computer-assisted instruction. Methods and technics from education, psychology, statistics, and computer program design are integrated into phases and functional steps in order to produce a general systematic logic for conducting research. Attention is given to the retention as well as the learning process in the research area. The methodology is seen as applicable to producing computer-based, standardized programmed instruction courses on a more solid scientific basis.

1.0 INTRODUCTION

The purpose of this paper is to define a systematic methodology applicable to computer-assisted instruction. The need for such a methodology results from the poor quality of research that has followed the rise of programmed instruction technology. The proliferation of devices, gadgets, and programs has been much faster than the logic of matters will allow. It seems appropriate, not to criticize and condemn, but to assist in further development of the nature of programmed instruction and the various techniques that are used with it.

Everyone knows that Pressey's* early devices were built initially to test more than to teach. The "items" that went into his machines were questions and possible answers to which a student had to select the correct one. If the initial selection was not correct, the student had to continue to try the other alternatives until he discovered the correct one. As a criterion of success, the student had to go through the complete set of questions and answer them all correctly at least once. Since his early "teaching machines" were tried out on children, Pressey always provided a reward when the child met the criterion. Food (candy) was used as the primary, positive reinforcement for a perfect score.

At once it became apparent that if a learner had the opportunity to discover each correct answer, he would eventually, by repetition, learn all the items. In a sense, then, Pressey's device was a rote memory training device which indicated to both the learner and instructor when the learner achieved a perfect score.

What is not so apparent, however, is the actual kinds of techniques involved. First, since the items in the machine were test items, mental test theory and item construction applied. Secondly, a reward was given for a perfect score, hence, learning

*Pressey, S. L., A Simple Apparatus Which Gives Tests and Scores-and Teaches. In School and Society, Vol. 23, No. 586, March 20, 1926

principles were used. Finally, insofar as the items were presented an item at a time according to increasing difficulty, some form of "programmed learning" was incipient as early as the 1930's.

At that time, Pressey did not construct and develop a systematic methodology for teaching and testing. The above description shows that a methodology was implicit, however, in terms of mental test theory, learning theory, and programmed learning techniques.

Twenty years later, Skinner re-introduced the concept of teaching machines and added the concept of programmed learning. Skinner's primary emphases were on the application of research performed in his laboratory with animals to human learning and on the development of machines or other devices to control behavior.

Since Skinner's research aimed only at developing predictable correlations between *stimuli and* responses, he did not relate his work to current behavioral or physiological theory. In the same way, when he developed his conceptual framework in programmed learning, Skinner focused on producing reliable responses to known stimuli. With his laboratory apparatus, Skinner controlled and measured behavior by appropriately recording signals to and from the animal. Similarly, his teaching devices were used to control and measure human behavior.

It was nearly ten years ago that Skinner* wrote his article entitled, Teaching Machines. This article pointed out the possibilities of "shaping" behavior in ways desirable to an experimentalist or teacher. Skinner's article, however, did not treat the problem of education and training from a systematic, functional viewpoint because Skinner had no explicit theory containing causal statements or explanation. Many of his followers

*Skinner, B. F., Teaching Machines, Science, Vol. 128, No. 3330, 24 Oct. 1958.

began to apply the new teaching technology in an almost random fashion and sometimes totally unrelated to previously accepted methods in psychology and education. Skinner probably did not intend such a movement to occur nor has he condoned it.

Other workers, since Skinner, have seen the need for a systematization of the technique. An example is the logic of mathetics, which is a technique for the strict application of stimulus-response relationships to learning, especially applicable to vocabulary building. Mathetics* systematizes Pavlovian, classical conditioning and functional instrumental conditioning and gives a central emphasis to reinforcement. Thus, the "chaining" of behavior becomes an important concept. For example: "cat-gato" might be paired together several times until the student saw them "associated" with each other. Then, the student will be given these two words together and he pronounces them after the instructor, who will say: "Good," or some similar phrase. *This segment of training* follows a Pavlovian paradigm. The next segment in the mathetics technique is to give the student only the word, "cat," and to expect him to say "gato." If the student says "gato," then the instructor says "Good" or "Correct." This arrangement provides immediate reinforcement for the correct answer. Now, if the student must learn: "The cat goes," the classical paradigm is repeated and is followed by the instrumental paradigm. For example, the pairs: "the-el," "cat-gato," and "goes-va" are formed. Next: "the cat-el gato" pair is formed. Then, "cat goes-gato va" is followed by the chain "El gato va."

The mathetics paradigm and the Skinnerian methods utilize the simpler discrimination and generalization responses for learning paired associates and forming sequential behavior ("chaining") by a systemization, as noted previously, of classical

*Gilbert, T.F., Mathetics, The Technology of Education, The Journal of Mathetics, Vol. 1, Jan. 1962.

and instrumental conditioning. Crowder* introduces a different paradigm which raises questions in education and training not actually dealt with anywhere in any formal sense by Crowder, since he gave considerable energy to the design of electronic and optical-mechanical devices more sophisticated than Pressey or Skinner's simple machines.

The terms "extrinsic" and "intrinsic" programming are often applied to programmed learning. These terms refer to the determiner of what next will be displayed to the learner. In Skinner's extrinsic programming, the response made by the student does not determine what he receives next. Since the student's behavior is being "shaped" in extrinsic programming, it matters little if his response is correct or incorrect because he will receive only the correct response next. The teaching device has complete control over the learner's behavior in the extrinsic programming technique whether the device be in a book or in a mechanical form.

In Crowder's intrinsic programming technique, the teaching device has less control over the learner's behavior. Organized information is presented and this is followed by a question with a set of alternatives in which only one in the set is correct. The learner must make a decision based on the information and the question and then select an alternative. If the teaching device is mechanical or electronic with an optical display, the learner presses keys indicated after the alternative. The depressed key produces the feedback page to the learner. If the teaching device is in book form, the alternative is keyed to a nonsequential page to which the student turns to determine if he is correct or incorrect. If correct, he receives new information, a new question, and a new set of alternatives. If incorrect, he

*Crowder, N. F., Automatic Tutoring by Intrinsic Programming. In Lumsdaine & Glaser's, Teaching Machines & Programmed Learning: National Education Association, Washington, D.C., 1960.

is given additional information on the nonsequential page in the form of hints as to why he is incorrect, but he is not given the answer yet. He is required to return and re-read the original information, question, and alternatives. The alternatives are written so as to determine whether the student has remembered previous information or perhaps has lacked prerequisites for the course in the first place. Sometimes a student is careless due to inattention; at other times only partial learning may have occurred.

It is at this point that one reason for a more systematic approach to programmed learning is needed, and later, such an approach will be described in detail. The technique of merely slicing information into words or paragraphs, and either asking the learner to reproduce the information or recognize it again when asked an appropriate question, certainly can not be the whole of teaching. Most educational and training institutions give at least casual thought to prerequisites, content, sequencing, examinations, and levels of achievement. Almost no one can continually pay attention, however, to the real-time behavior of the student as he learns information by book, teacher, or device.

Moreover, what is prepared as information by teacher, curriculum specialist, or textbook writer is like a game in which these educationalists predict that the information that they have organized is isomorphic with the way a student will learn. A closer scrutiny of examinations in classroom, lecture, or textbook reveals that little learning occurs compared with cost and efficiency. Textbooks and other books are often written for colleagues after the first chapter. Lectures often entertain or produce a quite unintended effect. Classroom discussions are often filled with irrelevancies.

Programmed learning tries to minimize these deficiencies and extravagancies but in so doing overlooks certain methods and technics that have scientific sanction. The Skinner method

desires that the correct answer be constructed by the student. The Crowder method desires the same thing but allows the student to recognize the correct answer among several alternatives.

The question arises, however, as to whether course objectives, course prerequisites, content sequences, programmed material, and examinations are methodologically consistent with each other in either the Skinner or Crowder programming techniques. Another question is: Does that which has been learned by a student have any reliability over time?

In several publications, Mager* has stressed behavioral objectives and the conditions under which the objectives should be specified. Too often courses are programmed or taught without an explicit statement as to what is to be expected of the student at the termination of instruction. Too often course objectives are stated in language that defies an instructor's ability to program or teach such objectives or an examiner's ability to test them.

Recently, Bloom** and his associates have made a determined attempt to classify behavioral objectives in terms of types of examination questions. These differ from Mager's concepts of behavioral objectives in that Bloom's objectives are actually examination questions classified according to six general principles: Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. Detailed sub-categories are subsumed under these six categories. For example, in the category, "Synthesis," the student is expected to "derive a set of abstract reactions," which is a sub-category. Mager does not reject this approach but he adds: That the conditions under which the behaviors are to be emitted should be operationally specified;

*Mager, R. F., Preparing Objectives for Programmed Instruction. San Francisco: Fearon Publishers, 1961.
**Bloom, B. S. (Editor), Taxonomy of Educational Objectives: Handbook I: Cognitive Domain. New York: David McKay, 1956.

that the student should know what is expected of him at the beginning of the course; and that this information be known to student, instructor, and examiner. For example, Mager probably states a course objective this way: "Given a linear algebraic equation with one unknown, the learner must be able to solve for the unknown without the aid of references, tables, or calculating devices." Bloom's* emphasis is on the examination itself and the kind of objective to be tested. For example:

> "2.00 Comprehension
>
> "2.10 Translation from symbolic form to another form, or vice versa.
>
> "18. Coulomb's Law of Electrostatic attraction states: 'The force of attraction or repulsion between two charged bodies is directly proportional to the product of the charges, and inversely proportional to the square of the distance between them.' If F is force, Q and Q' are charges, D is the dielectric constant and d is distance, a mathematical statement of the law is:
>
> a. $F = \frac{Q}{Dd^2}$. b. $F = \frac{QQ'}{D^2 d}$. c. $F = \frac{QQ'}{Dd^2}$.
>
> d. $D = \frac{Q^2}{Fd}$. e. $d = \frac{QQ'}{DF}$."

Essentially, Mager and Bloom agree. Mager's system does not add the examination question itself; however, the specification of the behavior to be achieved is explicit in both cases.

*Bloom, op. cit., page 104.

Mager* has also been concerned with content sequencing as have many curriculum specialists and textbook writers. He distinguishes between instructor-centered content sequencing and learner-centered sequencing. Instructor-centered content sequencing is characterized by some kind of logical order: temporal, spatial, inductive, deductive, general-specific, etc. With learner-centered content sequencing, the learner combines on an immediate need basis "what he needs to know with what he already knows**." But such an arrangement must be experimentally, rather than intuitively, determined as will be suggested later.

The terms "initial" and "intermediate" behavioral objectives have been proposed and analyzed by several writers in connection with content sequencing as well as with behavioral objectives in a terminal sense (examinations). In this connection the terms "mediating" and "transitional" behaviors are often used for "hints" and "prompts." Some writers have preferred that all of these terms be defined operationally by such devices as achievement tests. The distance between steps in difficulty of a population of test items would constitute the segments in the sequencing of content. It should be noted, however, that such tests are generally constructed by teachers or examiners. Cureton*** has pointed out that "the <u>de facto</u> aims of an educational program, and of every part thereof, consists of those acts on the basis of which the students and the program are in fact evaluated. If any stated aim is not analyzed into specific actions, and those actions observed and scored and reported, the statement is no more than empty verbiage." Thus, Cureton is pointing out the relationship between objectives, content, and measurement.

---

*Mager, R. F., <u>On The Sequencing of Instructional Content.</u> Palo Alto, Calif., Varian Associates, 1961.
**Mager, op. cit., page 412.
***Cureton, E. E., "Validity," Chapter 16 in <u>Educational Measurements</u>, E. F. Lindquist, Editor, Washington, D. C., American Council on Education, 1951.

Ebel* states, in general agreement with Cureton, that it is "... easier to define desirable behavior in terms of test exercises than in terms of curriculum procedures." Ebel would "... regard the hypothetical population of items from which [a] particular sample is presumed to have been selected, as constituting a better operational definition of the goals of achievement" than a single test sample.

Cureton** has defined the criteria for a population of test items for initial, intermediate, and terminal behaviors as being characterized as follows:

a. "The acts or operations of which a [function] is composed."
b. "The materials acted upon."
c. "The situations in which these acts or operations properly take place."
d. "The results or products of these acts or operations."
e. "The particular aspects or features of the acts or of their results or products which are to be considered as germane to the function."

Adkins*** points out in her definition of a test that it is a "... means of drawing inferences about persons, based upon their responses to a sampling of a field of behavior." If this definition is used for a test and if initial, intermediate, and terminal objectives can be stated in terms of test items, then the student's progress can be determined in a measurable way. Moreover, such a determination acts as a constraint envelope on content sequencing because it excludes the sequencing of irrelevant content. The original or parent population of test items

*Ebel, R. L., "Obtaining and Reporting Evidence on Content Validity." Educational and Psychological Measurement, XVI, 3, 1956, 269-282.
**Cureton, op. cit.
***Adkins, D. C., "Measurement in Relation to the Educational Process," Educational and Psychological Measurement, XVIII, 2, 221-240.

can be sampled progressively to determine the student's achievement.

Thus, it is seen that there is a relationship between content sequencing and the behaviors that must be learned. This relationship is a constraint on what is sequenced by the criterion of relevancy. However, Mager's finding that content-sequencing must be based on student needs rather than upon the arbitrary logic of an instructor must be discussed. What an individual learner needs to learn in order to achieve specified objectives depends upon very specific needs at the moment. Seldom is an instructor capable of fulfilling the range of such needs while conducting conventional classroom teaching. How, then, can the student-centered instruction concept be useful? Would a given programmed course have to be written at several levels of IQ, or would one segment be devoted to prerequisites and another to instructing terminal behaviors?

The answer can probably only be determined by experimentation as a course is pre-tested. A representative sample of students must be selected to pre-test the programmed course. In addition, a complete recording of their behavior while learning must be provided. The following is a partial list of information on such a recording:

a. Frame number
b. Time per frame
c. Number of errors per frame
d. Average time per frame for all students
e. Total number of errors per student
f. Total time to learn course to criterion
g. Number of frames used per student
h. Total number of frames
i. Real-time

Another question to which an adequate answer must be determined is the "prime path" for sequencing content. It has been

observed that students, who take the course, whether programmed or not, whether programmed linearly or non-linearly, or whether extrinsic or intrinsic will have differing needs. It has also been pointed out further that the path chosen by the instructor, by a conventional textbook, or by the usual programmed course is not the path of the individual student.

When content is sequenced linearly as in a Skinner type program, there is no "paths" except one--the teacher's. Linear programming is best seen as a special case of learning in which individual words, phrases, symbols, or forms must be acquired by the learner. Nonlinear programming is best seen as a more general case of learning in which problems are solved or concepts formed based on an organized unit of information, and there is a <u>choice</u> of alternatives from which the learner makes a decision. He is being instructed and evaluated almost simultaneously before he proceeds to the next organized unit.

Dressel* has pointed out the close connection between instruction and evaluation by stating that: "Evaluation does not differ from instruction in purposes, in methods, or in materials, and can be differentiated from instruction only when the primary purpose is passing judgment on the achievement of a student at the close of a period of instruction." He also emphasized that: "Testing for knowledge should be supplemented and even in part replaced by broad, pervasive, and continuing evaluation or assessment which becomes a major part of instruction and, therefore, indistinguishable from it."

Now the question arises frequently as to whether recognition or reproduction of the answer produces superior learning and retention, given that instruction and evaluation should be almost simultaneous and continuous. In Skinner's linear programming, the student tries to remember the paired words or phrases

---

*Dressel, P. L., "Evaluation as Instruction," in <u>Processing, 1953 Invitational Conference on Testing Problems</u>, Educational Testing Service, Princeton, New Jersey, 1954.

after they are presented and writes them down. The retention time is of the order of 0.5 seconds. His reproduction time may be of the order of one or two magnitudes greater than the immediate memory requirement. The "response time" plus the "memory time" results in a very short "buffer storage" and "retrieval" period. It is an empirical matter whether this produces the long-term storage both necessary and sufficient for later accurate recall. In the Crowder or nonlinear system of programming, information received by the student is organized into paragraph size more often than short sentence or paired associates size. A question follows this information rather than a blank space as in linear programming. Then, a set of alternatives is presented in nonlinear programming and these alternatives contain distractors as well as the correct alternative.

The combination of the question and the alternatives is called an "item." This combination is based more on understanding of the organization or sequences of thought of the material than on the memory for it. For example, the student should be able to recognize a paraphrase, a summary of what was given in the informatory paragraph, or a recall of a sequence of manipulations as in an algebraic equation. Word meanings are emphasized in nonlinear programming more than word pairs or short sequences as in linear programming.

There is nothing in linear or nonlinear programming which prohibits a student from repeating given information until he memorizes it. In nonlinear programming, rote memory is seldom used since the emphasis is on the meaning of organized material. This is not to say that linear material is unorganized because the sentences, even phrases, are organized. Information, however, is fragmentary. It is often difficult "to get the picture" or general framework of an idea or concept. In nonlinear programming, the Gestalt of an idea or sequence is presented in its entirety as a unit--it is organized for the student.

The question arises as to the amount and kind of practice necessary in the two types of programming. Should practice immediately follow initial learning or should "periodic review" be inserted in the program? How many times should the student practice the whole programmed course in order to achieve a given criterion for practice? Does overlearning help maintain longer recall? Again, as for other problems of human learning, the distribution and the amount of practice questions must be matters of research. Different distributions of practice must be tried along with different amounts of practice.

Nonlinear programming is often called "branching programming." In this type of programming, one path, the prime path, consists of the set of correct alternatives to all the frames. Each frame has only one correct alternative among the available alternatives. Now it occurs that if the alternatives have been designed properly, at least one <u>wrong</u> alternative should test whether a student has the prerequisites or has learned material previous to the present frame. Another <u>wrong</u> alternative should test whether the student has made tacit assumptions which lead to errors or irrelevancies in deduction or induction, etc. The remaining alternatives should test for understanding concepts or following correct sequences, except that one of these alternatives should be wrong and contain a careless, highly probable error. These alternatives should provide some difficulty in discrimination between them and one should be based on a lucid understanding of the concept or problem.

The first wrong alternative should lead to extensive "branching." A student should be directed back to a series of frames other than the one on which he made his error, but the "washback" should lead back finally to that frame which caused him to make the serious error. This washback provision enables learners, at different levels of IQ or aptitude, to progress at different rates. It should not be expected in all cases that the learner, once through a program, will have retained much of

what he has learned even for immediate recall and that he will have saved himself from practice because he has learned so rapidly. It is a matter of research to find out how much practice is necessary within a programmed course as well as how many times the total course should be repeated to reach the practice criterion, especially if the criterion should be, say, two perfect practice sessions beyond the first perfect time through the entire course. The results of such research would provide excellent information to the programmer as to where additional material for practice should be inserted even for the most talented student. A brilliant student may go through the prime path without error, but upon examination, some forgetting will have occurred or he may have guessed and obtained the correct answer during practice and then fail the item on test.

In the case of the lower aptitude and IQ learners, branching to easier and/or review material enables these learners to re-learn not only prerequisite material, but more importantly this provision for having discriminating alternatives for each item in nonlinear programming further enables the programmer to integrate his techniques with mental test techniques.

In the case where branching enables the lower IQ and lower aptitude learners to depart from the prime path, and at the same time have the opportunity to return to it, nonlinear programming allows for continuation of learning in the face of actual difficulty with a frame, yet reduces any undesirable negative affect and resultant perseveration of error or regression from the learning situation. Presumably in the linear programming techniques, "steps" are so small--word pairs, small phrases, hints, word meanings--that error probability is very low.

The question arises as to the utility of error in learning and performance. Seldom does the real world organize information in such a way that the learner's receptors and brain needs make little effort to use the information provided for his

own benefit. More likely an alternative is selected from minimal or overwhelming information, all of which is loosely organized. Such is the set-up between Man and Nature, between Man and Man, and within Man himself. Error is used to re-direct the individual, but in Man's relation to Nature, seldom does Nature tell him what to do next. Similarly, interactions between men many be extensive or frequent before relevant information is received and the required behavior occurs.

Error can be used to improve a programmed course during experimental trials before release to the user.

The significance of error is at what point in time it is made. It is important, therefore, to have a real-time record of error; what alternative was selected first, second, third, etc. Of course, the count of the number of errors per alternative for item analysis must be made, but this statistic is not as important for the learner or the programmer as is the selection of a wrong alternative the first time. It is important for the learner to get the information immediately. It is important for the programmer because he may need to revise the prime path, insert more branching, increase practice problems, or perhaps to combine a number of these factors.

The discussion above about error is relevant to non-linear (branching) programming. In the case of Skinnerian, or linear programming, error is not supposed to occur because what is to be learned is isolated and is presented, spatially and temporally contiguous, with already-learned material of very short length. If the student's reproduction is wrong or right, during learning, he is given the correct answer immediately. The programmer assumes that his sequencing as well as his programming is correct in that the student can go from sequence to sequence and frame to frame without error.

If the student could be asked to make a prediction as to the probability that the alternative he has decided upon is correct, then the subjective sizes of the "steps" can at least be

partially evaluated from the student point of view. This would give Mager's concept of student-centered sequencing a more adequate operational test. It also would provide a basis for an operational comparison between the prediction of the next step by the programmer and the prediction of the next step by the learner.

Now, since in original programming, the programmer predicts that the next "step" is in the proper sequence, it follows that

(1) $$p'_P = 1.$$

If a student also predicts that he would be correct with a probability of 1, then

(2) $$p_S = 1,$$

and

$$p'_P - p_S = 0.$$

The question arises as what to expect if $.00 \leq p_S \leq 1.00$.

This situation may look like this for a single student for a single frame (where a student picks alternative "b" and assigns a "p" of .75 to it indicating he is quite sure he is correct):

(3)

a. $p_S = 0$

b. $p_S = .75$

*c. $p_S = 0$; however, $p'_P = 1$

d. $p_S = 0$

*correct alternative

But for the same frame, the distribution may in early training look like that shown in Matrix [1] for 30 experimental students: (The matrices which follow are hypothetical illustrations).

[1]

	PROBABILITY					
i	.00	.25	.50	.75	1.00	n_2
a.	0	0	1	1	0	2
b.	0	1	1	1	1	4
*c.	5	3	2	8	3	21
d.	1	0	1	0	1	3
n_1	6	4	5	10	5	30 = N

It should be kept in mind that the focus is on experimental students using them to pre-test the program prior to formal training with it. Most, if not all, programs are usually pre-tested by post-training examinations and after-the-fact statistics are used for analysis.

It can be seen by Matrix [1] that 70% of the students chose the correct answer and only 10% had 100% confidence in their choice. Some students had high confidence in wrong choices.

What is the meaning of these other distributions? Experimental data on matching** in which the student tends to "track" that which is rewarded most frequently might apply here, in the case of practice on the whole program. Then an item would be repeated and Matrix [1] would appear by, say, five repetitions for all students as follows in Matrix [2] :

*Correct alternative, c ; n_1 = column sums; n_2 = row sums; N = $n_1 + n_2$.
**Koch, S. (Editor). Psychology: A Study of a Science, Volume 2, page 413. McGraw-Hill, 1959, New York.

[2]

	PROBABILITY					
i	.00	.25	.50	.75	1.00	n_2
a.	0	0	0	1	0	1
b.	0	0	1	0	0	1
c.	0	0	0	0	27	27
d.	0	1	0	0	0	1
n_1	0	1	1	1	27	30 = N

Now 90% of the students have learned by practice to get the correct answer and have confidence in it.

Then the question arises as to the implications of continual distribution of errors and continual lower subjective probability, p_S, after practice, <u>to the programmer</u>.

In the example of Matrix [2], confidence is rather high in two cases of error. In the remaining case, error persists and confidence may be low even after, say, five practice sessions. It seems, logically, that subjective probability may persist, decrease, or increase while a student shifts his response from error to error, from correct response to error, or from error to correct response. Thus, a five-trial <u>practice</u> Matrix might look like this:

[3]

	PROBABILITY					ALTERNATIVES			
t	.00	.25	.50	.75	1.00	a	b	c*	d
1.	-	-	a	-	-	1	0	0	0
2.	-	c	-	-	-	0	0	1	0
3.	-	-	a	-	-	1	0	0	0
4.	-	-	-	-	c	0	0	1	0
5.	-	-	-	-	c	0	0	1	0
n_1	0	1	2	0	2	2	0	3	0

*c = correct alternative; t = trial number

This Matrix illustrates an error followed by learning to criterion (two successive) of the correct response. However, the student appears to have more confidence in his guessing ability than in obtaining the correct answer since he perseverates on the "a" answer on the third trial after getting it correct once. Theoretically, it could be speculated that the "a" alternative is a stronger association or "fits-in" with some previous organization of learned material in the early trials.

One implication of this for the programmer is to correlate the above analysis with IQ level and to derive actual content sequence and branches from the prime path. This procedure would reduce the programmer bias and allow for student "needs to know" to emerge. For example, a matrix which shows both an increase in confidence or predicted probability of correct answer with practice trials as well as an increase in probability of learning the correct answer with practice trials would mean no branching is needed--only practice--if 100% of the learners got it correct after two practice trials in an experimental test of the program.

If 90% of the learners got the correct answer twice in succession on a frame in two trials then no branching is needed on that item frame.

If 90% of the learners took more than two trials to obtain criterion (two correct trials in succession) then branching is indicated based on the alternatives for that frame. Care should be taken that an item analysis of alternatives shows them not to be non-discriminatory. For example, if alternatives or the questions are poorly written, two right answers, no right answers, ambiguity, etc., may arise.

If 90% of the learners took more than five trials to reach criterion then the content sequencing must be examined from the point of view of the adjacent frame statistics. The possibility exists that adjacent frames have not been sequenced properly or the "steps" are too large. If the adjacent frames show no

excessive trials to criterion then the internal characteristics of the frame must be examined to see if the presentation of the information, the formulation of the question, and the statement of the alternatives are free from fault.

The question arises as to what use should be made of the two types of frames which provided the learner with information about his selection of alternatives. One generally informs him that: He has selected a certain alternative, that the alternative he selected was wrong, and that he has arrived at the wrong answer because of carelessness about a detail, lack of understanding of the teaching frame, lack of discrimination among the alternatives available to him, failure to have learned or remembered information from a previous frame, or failure to have met all prerequisites for the course. More than one of these reasons may have combined to cause his errors. There may be other reasons why he might make errors, but the concern here is not with error analysis.

The other type of frame informs the learner that: He has selected a certain alternative, that the alternative he has selected was correct, and the reasoning by which he _should_ have arrived at the correct alternative. On this same frame is included the next step in the program, although a separate frame can be used for this purpose.

The first type of frame is called the "wrong answer frame." The second type is called the "right answer frame." It is conceivable that a third type of frame might be called the "new information frame," whenever the right answer and the new information frames are not combined.

Returning to the question of the utility of the "feedback" frames (right answer frame and wrong answer frame), it appears that additional information should be provided the learner in the form of reinforcement, both positive (gain) and negative (loss). A point system might be a sensible approach to reinforcement. The goal of the learner may be the criterion of 90% correct responses twice in succession on the total programmed course. If

there are 100 "steps" in the prime path, the student could fail any ten of them on his first practice session and meet the criterion. On his second practice session he could still fail any ten items. It must be noted, however, that each wrong alternative subtracts one point; hence, there is more than one possibility to subtract points and only one to add points where a system of reinforcement is used.

In order to achieve a criterion of learning using 90% correct twice in succession during practice, points should be added or subtracted in such a way that one point is added for each correct alternative selected and one point subtracted for each wrong alternative selected. When testing the program with a computer, the learner can observe his cummulative "rights" and "wrongs" as he responds (as well as information about each frame he has selected). For example, if there were 100 correct frames, the student would be allowed only ten wrong selections twice in succession during practice to achieve a criterion of 90% correct twice in succession.

The point system of reinforcement and the criterion level set in relation to it allows for consistency between them in the nonlinear programming technique. In the linear programming technique, the system proposed herein does not appear to be applicable. If the student should select a synonym or some other approximation as his response, it would be entirely a subjective matter to assign points for correctness in the linear system. For example, a student might write "reflects" for "reflex."

Another question arises concerning the reliability, and hence, the cost of education and training by means of programmed learning. The subject of cost is not taken up in detail in this analysis. If a student has learned a given amount of material and if he has achieved the minimum criterion of 90% twice in succession during practice, and if he also achieves the minimum criterion of 90% on a test of the material after a fixed interval of time, then it might be stated that the training was highly reliable because the criterion was met without excessive training time. The cost,

however, may have been high to achieve the criterion in terms of administrative expense.

Additional costs are incurred when forgetting occurs. Some forgetting will occur even for short periods and some remembering will occur even for long periods. The problem of skill and knowledge maintenance must be dealt with in order to minimize losses when skills and knowledge are infrequently used. One technique used to minimize this future loss is overlearning. But overlearning increases costs if carried beyond the point where increasing marginal increments of practice produces decreasing marginal increments of recall. For example: If, after 100% overlearning, the score one achieves is only one or two points higher than 50% overlearning, then 50% overlearning is more economical for most situations. What would this imply for programmed learning courses? Should practice be massed after each step or should the whole course be practiced, or should practice in the form of recurring review within the course be constructed? These are empirical matters and only longitudinal studies can answer them.

A final question remains and that is the suitability of present mathematical models to programmed learning. As explanatory models, mathematical models have little utility. As predictive models based on very simple explanatory models, mathematical models have some utility for paired associates learning. Whether or not an adequate explanatory model and/or mathematical model is available to assist in explaining and predicting problem solving type learning is not the subject of this paper. When appropriate systemization has been made of empirical matters, then it is probable that resulting data may be fitted by various types of mathematical models. Adequate explanatory theories will probably await the empirical solutions and mathematical representations.

The remainder of this paper will present a functional and systematic approach to matters discussed in this introduction.

2.0 THE SYSTEM APPROACH TO COMPUTER-ASSISTED INSTRUCTION

This section will discuss in terms of the ideas and techniques of the previous section an integrated, systematic, functional methodology for computer-assisted instruction. The approach to be discussed is not directly concerned with terminal devices linked to a single computer nor with networks of communication hardware designed to service large educational systems. It is not concerned with types of displays, controls, consoles, nor computer programs designed to interface with input-output devices and the learner. Finally, it is not concerned with particular course content or types of educational systems. It is recognized, however, that the production of learning materials may require a wide range of selection of hardware and software for various purposes. Some educational productions may be so elaborate as to require talents, techniques, and hardware of the motion picture industry. Other educational productions may require only programmed texts with no more additional and elaborate talent than a good illustrator, draftsman, or artist. Still others may require all of these technics plus general managerial, engineering, and production talents of the broadcasting industries.

It is the attempt of this paper to integrate certain educational and psychological developments that have been accepted through research and theory in these scientific fields. Parallel to these developments has been the increase in engineering technological breakthroughs, which when integrated with scientific disciplines, provide mutual benefits for both. The demand for the training and instruction of technicians, scientists, and humanists in both the underdeveloped and developed nations has operated to increase the supply of personnel, hardware, and software to meet this demand; however, systematization of these various disciplines has barely started.

The reader should refer frequently to Fig. 2-1, Functional Flow Diagram: Plan for the Development of a Systematic Methodology for Computer-Assisted Instruction (CAI), in order to obtain a general overview of the developmental plan and detailed functions to be performed. These are eight phases of development and 24 functions to be performed during the eight phases. The phases and functions will be discussed in the following paragraphs:

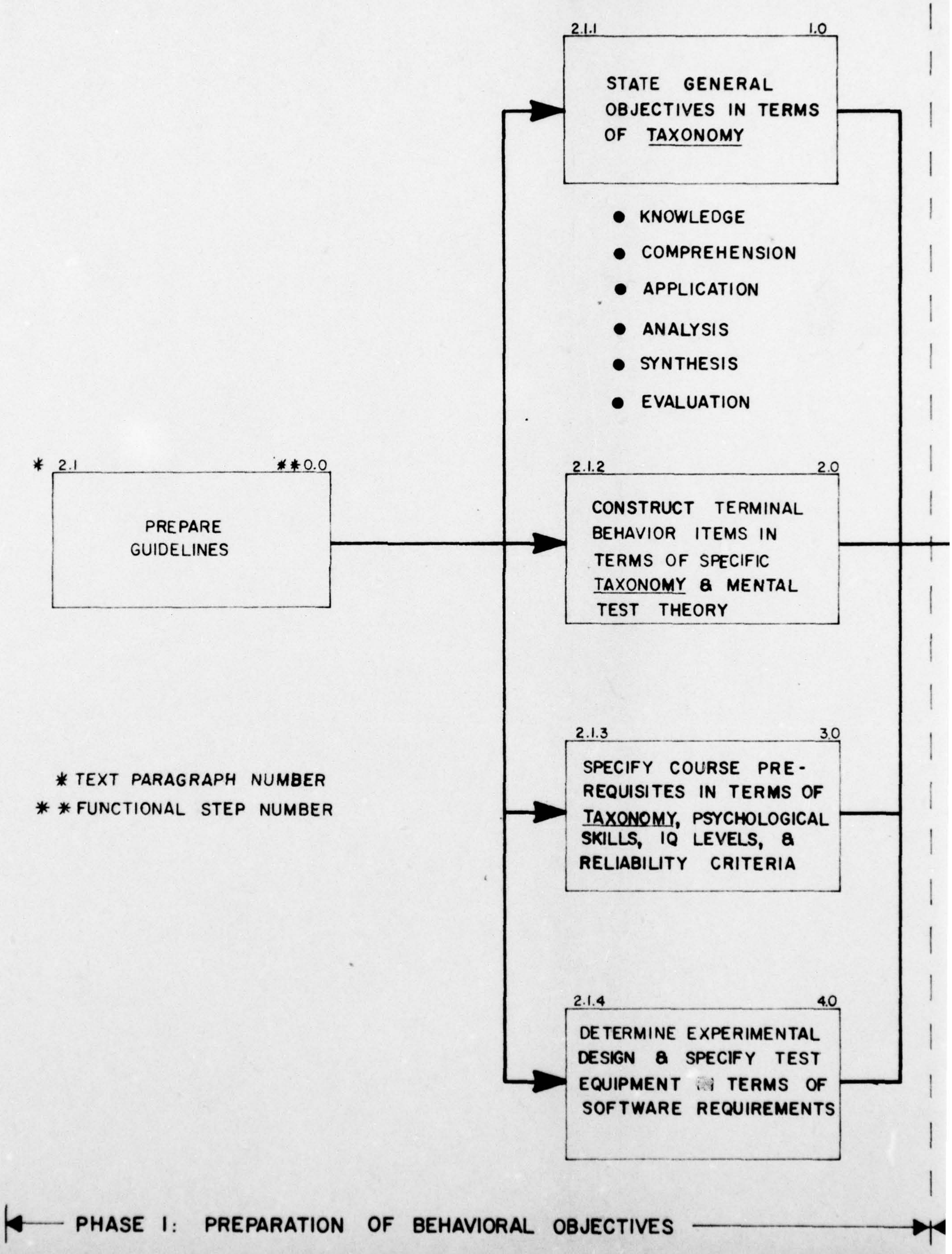
2.1 Phase I: Preparation of Behavioral Objectives

This phase begins when Guidelines (0.0)* are received in written form from the agency authorized to specify the requirements for a course of study and to direct its preparation under certain constraints. Within the constraint envelope specified by the establishing authority, general objectives are stated in terms of a taxonomy of objectives. Next, terminal objectives are constructed in terms of the stated general objectives, the taxonomy, and mental test theory. The third step is to specify course prerequisites in terms of the taxonomy, IQ levels, and reliability criteria. Since all courses using CAI (Computer-Assisted Instruction) must go through an experimental test, an experimental design is prepared during this phase to allow sufficient time for planning.

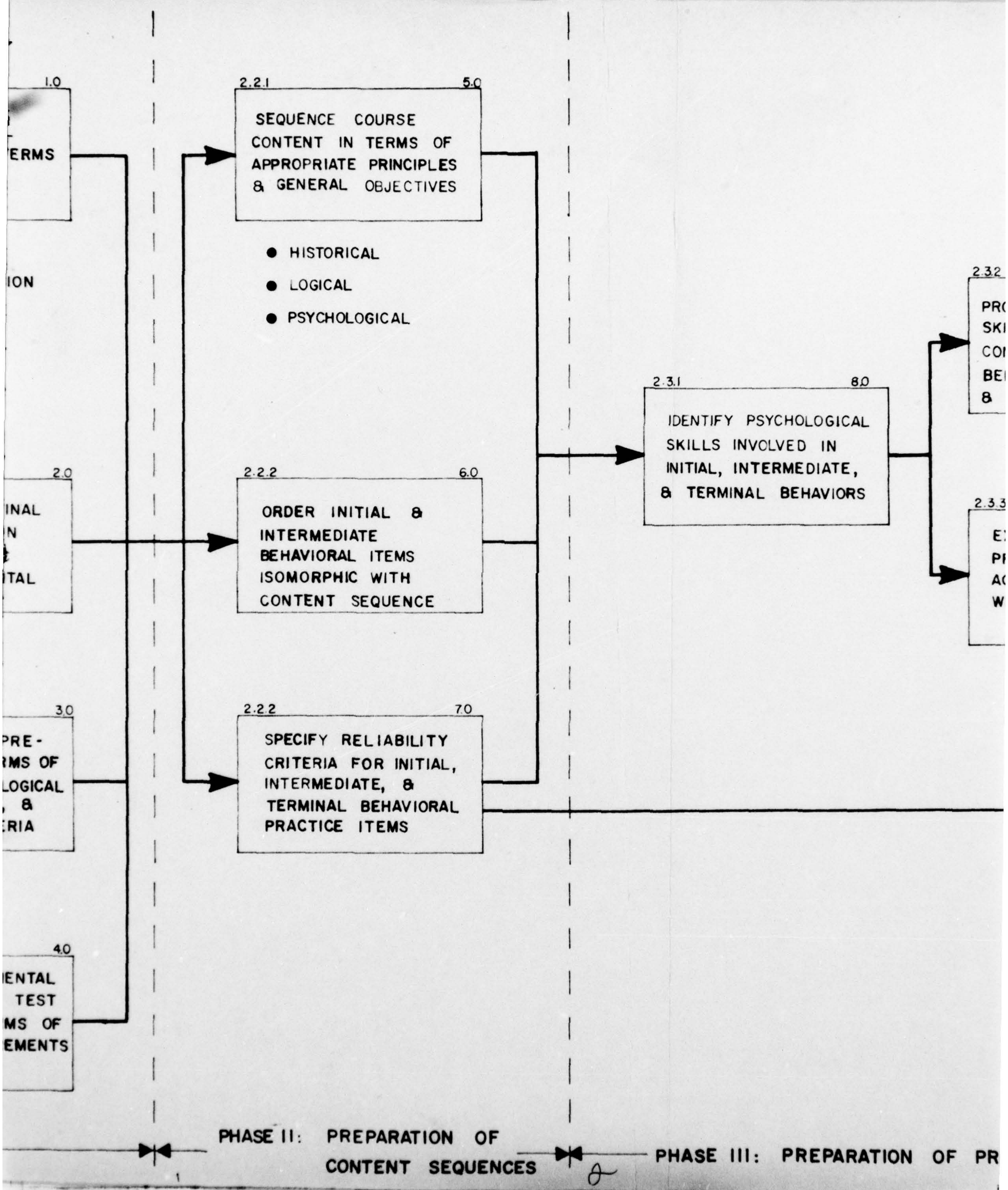
2.1.1 Statement of General Objectives (1.0)

General objectives are stated in such a way that all concerned are aware that the student can perform in certain specifiable ways at the termination of training. In order to do this, Bloom's Taxonomy** should be used. The general categories he uses are: Knowledge, Comprehension, Application, Analyses, Synthesis, and Evaluation. These general categories are broken down into more specific categories. The general objectives are classified in terms of these categories, using only those relevant for the behavior desired, in advance, before examinations are

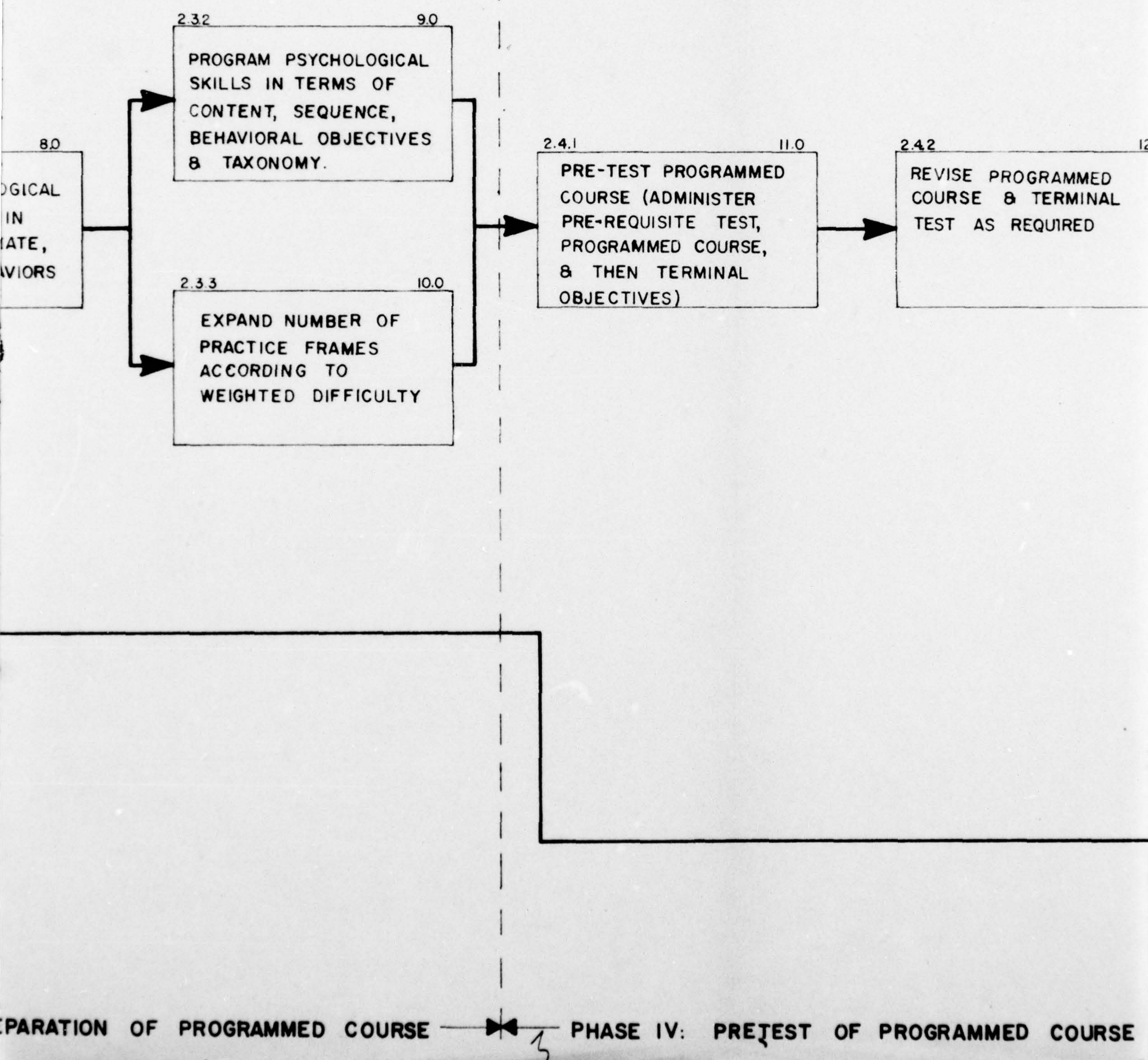
*This number, (0.0), refers to one of the boxes in Fig. 2-1.
**Bloom, B., Taxonomy of Educational Objectives, Handbook I: Cognitive Domain. New York: David McKay, 1956.



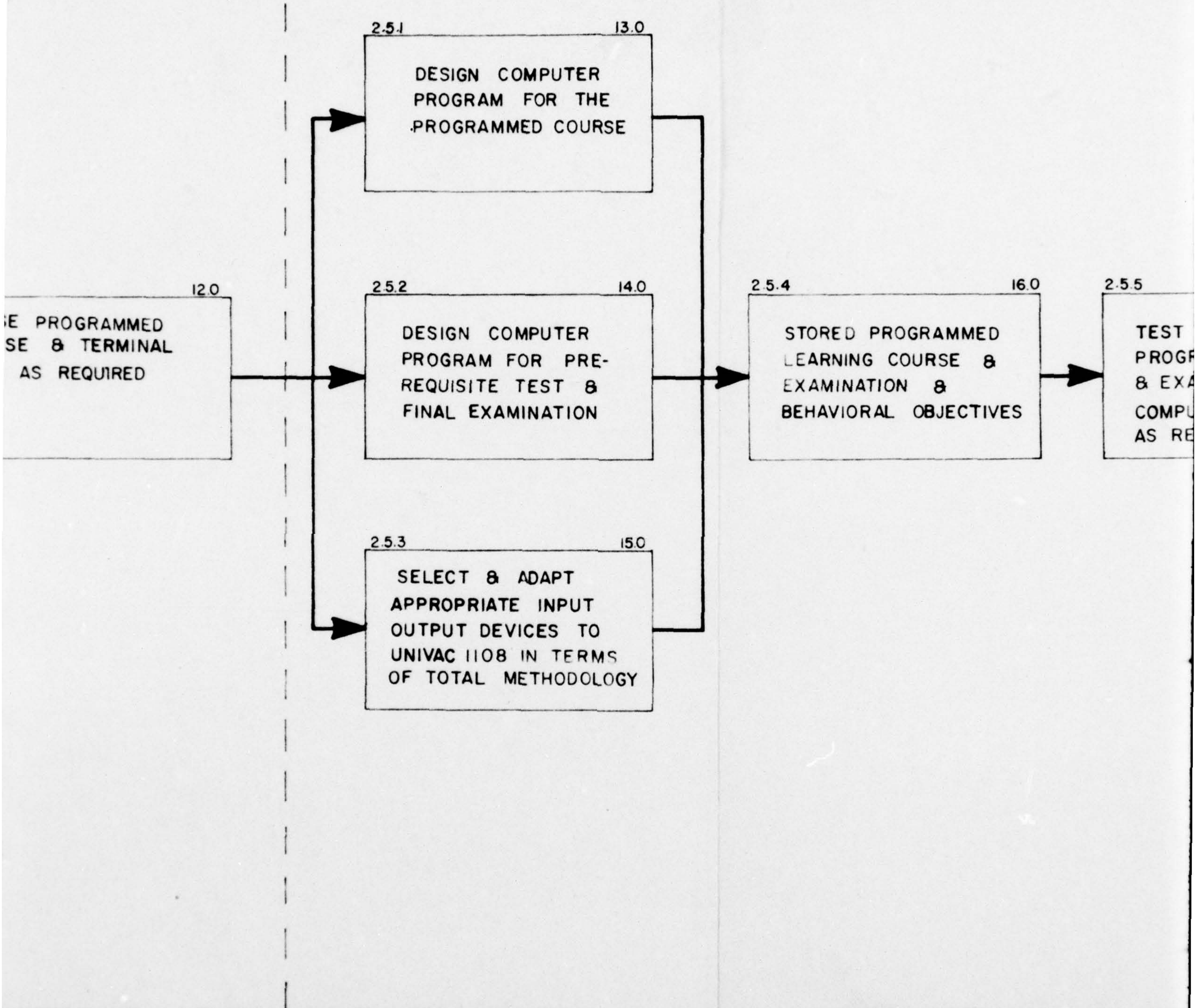
2.1.1
1.0
STATE GENERAL OBJECTIVES IN TERMS OF TAXONOMY
● KNOWLEDGE
● COMPREHENSION
● APPLICATION
● ANALYSIS
● SYNTHESIS
● EVALUATION
* 2.1
** 0.0
PREPARE GUIDELINES
2.1.2
2.0
CONSTRUCT TERMINAL BEHAVIOR ITEMS IN TERMS OF SPECIFIC TAXONOMY & MENTAL TEST THEORY
2.1.3
3.0
SPECIFY COURSE PRE-REQUISITES IN TERMS OF TAXONOMY, PSYCHOLOGICAL SKILLS, IQ LEVELS, & RELIABILITY CRITERIA
* TEXT PARAGRAPH NUMBER
** FUNCTIONAL STEP NUMBER
2.1.4
4.0
DETERMINE EXPERIMENTAL DESIGN & SPECIFY TEST EQUIPMENT TERMS OF SOFTWARE REQUIREMENTS
PHASE I: PREPARATION OF BEHAVIORAL OBJECTIVES


1.0
ERMS
ION
2.0
INAL
N
ITAL
3.0
PRE-
RMS OF
LOGICAL
, &
ERIA
4.0
IENTAL
TEST
MS OF
EMENTS
2.2.1
5.0
SEQUENCE COURSE CONTENT IN TERMS OF APPROPRIATE PRINCIPLES & GENERAL OBJECTIVES
● HISTORICAL
● LOGICAL
● PSYCHOLOGICAL
2.2.2
6.0
ORDER INITIAL & INTERMEDIATE BEHAVIORAL ITEMS ISOMORPHIC WITH CONTENT SEQUENCE
2.2.2
7.0
SPECIFY RELIABILITY CRITERIA FOR INITIAL, INTERMEDIATE, & TERMINAL BEHAVIORAL PRACTICE ITEMS
2.3.1
8.0
IDENTIFY PSYCHOLOGICAL SKILLS INVOLVED IN INITIAL, INTERMEDIATE, & TERMINAL BEHAVIORS
2.3.2
2.3.3
PHASE II: PREPARATION OF CONTENT SEQUENCES
PHASE III: PREPARATION OF PR

FUNCTIONAL FLOW DIAGRAM: PLAN FOR THE DEVEL

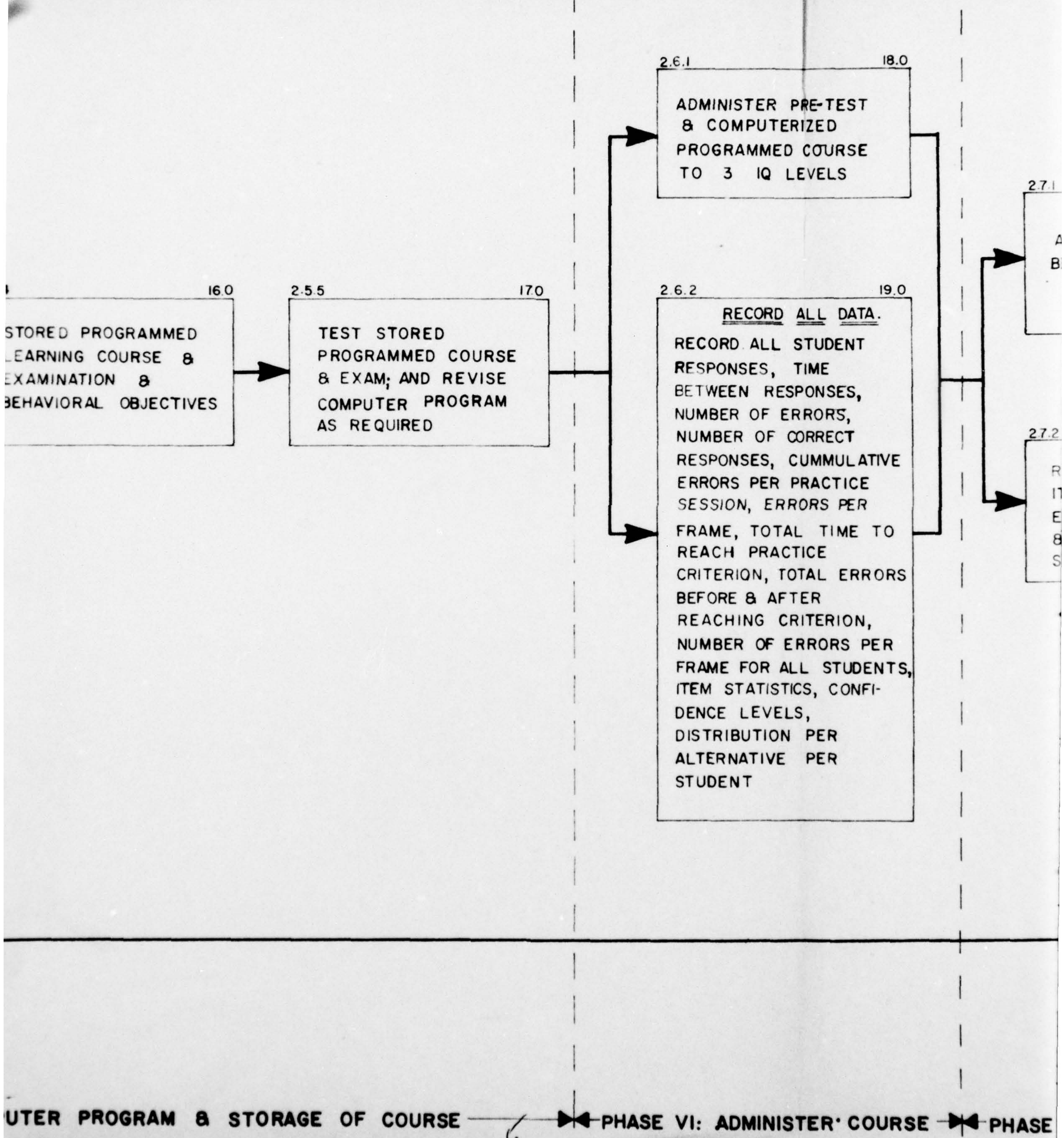


THE DEVELOPMENT OF A SYSTEMATIC METHODOLOGY FOR CO

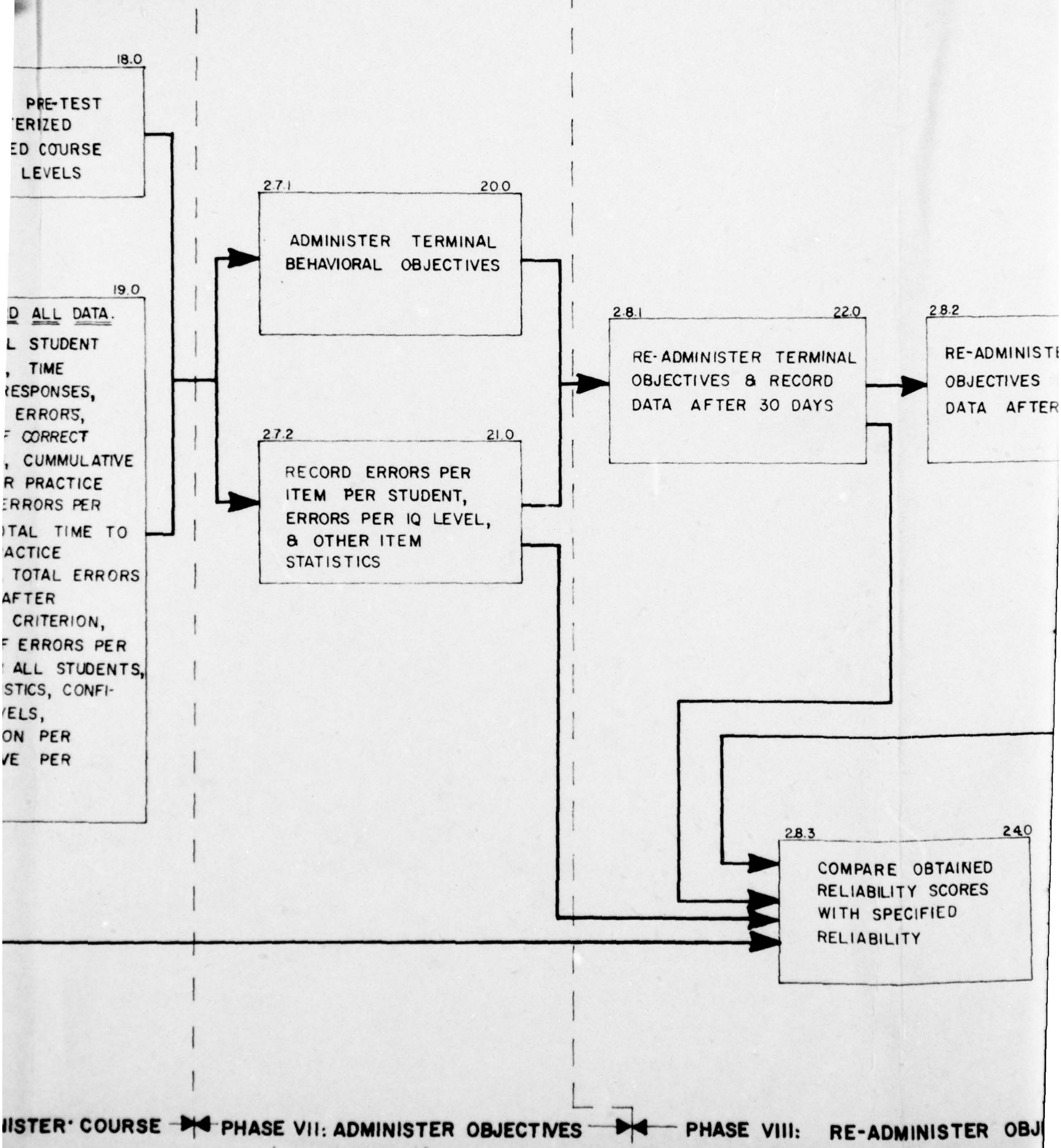


AMMED COURSE → ← PHASE V: PREPARATION OF COMPUTER PROGRAM & STORAGE OF

ETHODOLOGY FOR COMPUTER-ASSISTED INSTRUCTION (CAI)



RUCTION (CAI)


18.0
PRE-TEST
TERIZED
ED COURSE
LEVELS
19.0
D ALL DATA.
L STUDENT
, TIME
RESPONSES,
ERRORS,
F CORRECT
, CUMMULATIVE
R PRACTICE
ERRORS PER
TAL TIME TO
ACTICE
TOTAL ERRORS
AFTER
CRITERION,
F ERRORS PER
ALL STUDENTS,
STICS, CONFI-
VELS,
ON PER
VE PER
2.7.1
20.0
ADMINISTER TERMINAL BEHAVIORAL OBJECTIVES
2.7.2
21.0
RECORD ERRORS PER ITEM PER STUDENT, ERRORS PER IQ LEVEL, & OTHER ITEM STATISTICS
2.8.1
22.0
RE-ADMINISTER TERMINAL OBJECTIVES & RECORD DATA AFTER 30 DAYS
2.8.2
RE-ADMINISTE
OBJECTIVES
DATA AFTER
2.8.3
24.0
COMPARE OBTAINED RELIABILITY SCORES WITH SPECIFIED RELIABILITY
NISTER COURSE
PHASE VII: ADMINISTER OBJECTIVES
PHASE VIII: RE-ADMINISTER OBJ

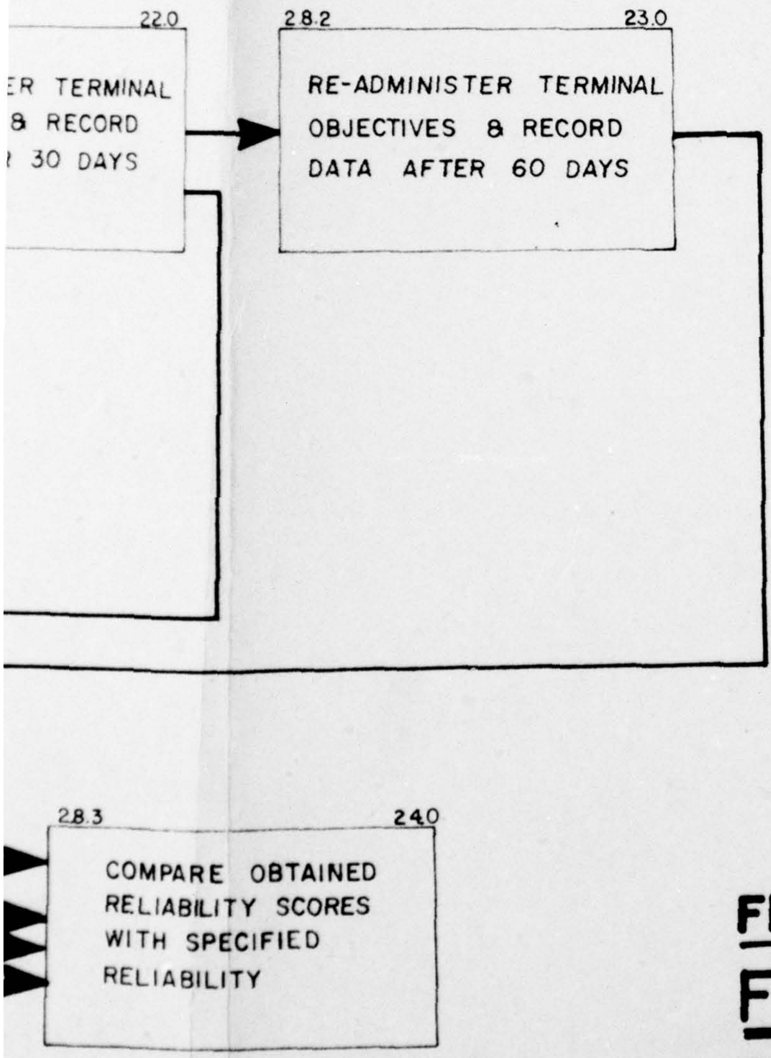


FIGURE 2-1:

FUNCTIONAL FLOW DIAGRAM

FJB/DC 1 JULY 67

P. 25

7

VIII: RE-ADMINISTER OBJECTIVES & COMPARE RESULTS

prepared, and communicated to all concerned with preparing the course, including the student. In addition, the conditions under which these performances are required must be stated. For example, if "Knowledge of Specific Facts" are required, then one must state whether they are to be given in sequence or some other logic. Or again, if "Application" is required, then one must state whether a module is to be repaired without or with a schematic. Many examples could be given from verbal and manual skills to be learned to illustrate the use of the taxonomy in specifying the general objectives.

2.1.2 Construction of Terminal Behavioral Items (2.0)

From the statement of General Objectives, terminal behavior items are constructed. Note that there is no directive to plunge immediately into content sequencing or programming at this time. In one sense, the terminal objectives (test items) are derived from the general objectives stated in Function 1.0, mainly because these latter objectives operate as constraints upon the test items. This step is taken to prevent irrelevant behavior from being taught or tested. For example, here is a test item selected from Bloom's Taxonomy*:

"2.00 Comprehension

"2.10 Translation from one level of abstraction to another.

1. A group of examiners is engaged in the production of a taxonomy of educational objectives. In ordinary English, what are these persons doing?

   A - Evaluating the progress of education
   B - Classifying teaching goals
   C - Preparing a curriculum
   D - Constructing learning exercises"

*Bloom, B. S., op. cit., page 99.

Each test item would be classified in terms of the Bloom Taxonomy according to both its general and specific categories. Items can be replicated to test the reliability of the learner's performance on the first examination as well as to test maintenance of the skill or knowledge through time by subsequent examinations. All the items together will be the population from which samples can be drawn for various purposes such as initial learning, practice or review, and final examinations periodically administered.

All items constructed should follow good test construction practices such as those given by Gulliksen* or Guilford** as well as others. In general, the construction of tests and the analysis and interpretation of their results require, according to Gulliksen***, the solution to "five major types of problems:

1. Writing and selecting the test items.
2. Assigning a score to each person.
3. Determining the accuracy of test scores.
4. Determining the predictive value of the test scores.
5. Comparing the results with those obtained using other tests or other groups of subjects."

It is not the purpose of this paper to describe this well-known area of psychology. It is, however, important that all behavioral objectives be constructed according to accepted practices.

2.1.3 Specify Course Prerequisites (3.0)

It is generally too early to select prerequisites for a course until after the psychological skills have been identified

*Gulliksen, H., Theory of Mental Tests. New York: Wiley, 1950.
**Guilford, J. P., Fundamental Statistics in Psychology and Education. New York: McGraw-Hill, 1950.
***Gulliksen, H., op. cit., page 2.

in relation to the behavioral objectives for the course. If this is done, the IQ levels, aptitudes, and reliability criteria can be based on specified requirements. For example, a prerequisite statement might read: Student must have an IQ of 100 as measured by the Otis Intelligence Test and be at the 95th percentile on the Detroit Mechanical Aptitude Test. If additional skills are required, then they, as well as intelligence and aptitude, must be specified in some numerical way in order to assure that these prerequisites are in alignment with initial behaviors required at the beginning of a course and not the terminal objectives. For example, a student may have to operate at the end of a course with exponents to a degree less than zero, but initially he must know how to operate with an exponent of "zero" or higher. Since some prerequisites may not be stated in IQ or aptitude form using percentiles, it may be necessary to state percentage achieved on tests using exponents, e.g., must have a score of 98% on examination, involving integer exponents, taken within one year prior to starting the course.

When initial behaviors are selected, they may very well include a major portion of the prerequisites, except for IQ and aptitude type items which would be irrelevant to training after it has begun.

2.1.4 Determine Experimental Design and Software Requirements (4.0)

In testing a CAI programmed course, it is necessary to specify in detail the experimental logic to be followed so that practical decisions based on the quantitative and qualitative parameters of the course can be made. Essentially this step consists of identifying all parameters and forming a matrix of their possibilities in order not to omit any parameters of interest. It is important to notice that the experimental emphasis in this paper is on the instrument being designed: The programmed course, not the comparison, say, of programmed instruction versus some other form of instruction, nor a comparison of linear versus nonlinear programming, etc.

The computer and appropriate input-output devices are the test equipment used to measure the parameters identified. The test vehicle is the programmed course. Since the course itself and its parameters of interest must be stored in a computer as well as the ongoing behavior of the student while learning, there must be a computer program written to interface between the programmed course, the student, and the computer itself.

It is during Phase I, however, that the computer programmer analyst must be given all information required in order that he can determine the inputs, outputs, storage, and retrieval data required for the computer operation.

2.2 Phase II: Preparation of Content Sequences

During this phase, content sequencing will be performed and initial, intermediate, and terminal items will be ordered in consonance with the content sequence established. Reliability criteria for passing these items during practice will also be established. The following paragraphs describe the details of these functional steps.

2.2.1 Sequencing of Course Content (5.0)

The initial content sequencing by the programmer will be historical, logical, and/or psychological. It is not known at this stage whether students' various needs would require a different order from that laid down by the programmer as was discussed in Section 1.0 of this paper. If the element of time, as for historical courses, is paramount throughout a course, then concepts and events should be ordered temporally to indicate their simultaneity and sequence. If explanatory processes are important, then their logic should order the content. If psychological principles are used, then such ordering as simple-complex, specific-general, discrimination-generalization, etc., can be used. It would be wise to consult a handbook* which states the facts and

*Blaisdell, F. J., Handbook of Human Learning: Applications to Education and Training. TRACOR, Inc., Austin, Texas, 1967.

principles of human learning in convenient form for ready use while content sequencing. The programmer will find that he will probably use all three at different times as he orders course content. It must be remembered that the programmer's initial version of the course content sequence is a pre-experimental version. The final content sequence is empirically determined.

The proper guide to use in sequencing course content prior to experimentation is to identify each terminal behavioral test item according to the principle chosen. Some items may require more than one principle; however, the programmer and subject matter expert should determine which principle should be introduced first in a particular sequence. The idea here is to have a continued isomorphic relationship between the sequencing of content and the general objectives of a course. This provision prevents irrelevancies from arising in courses and prevents relevant materials from being omitted where required. Too often what should have been taught in class is examined on a test or vice versa. For example, "all A's are B's" may have been taught, but the student may be tested on its converse which may never have been mentioned in class through neglect. Or another example, the validity of "all B's are A's" may have been taught while the meaning of "amphibolous" may not have been taught but will have been tested. This last example shows the testing of an irrelevancy but teaching a relevant concept. Such a procedure might even fail the brightest student.

If the terminal objectives have been written in terms of the taxonomy, then the introduction of irrelevancies on a test and the omission of relevant materials during instruction will have been minimized.

The state-of-the-art of concept sequencing does not promulgate rules for different kinds of subject matters. The subject matter expert and programmer must design the course together unless they happen to be the same individual.

A final warning for content sequencing should be expressed: WHAT must be taught and tested should not be confused with HOW something is to be taught and tested. That is why the historical, psychological, or logical principles must shift throughout a programmed course. Each of these principles may be used on different levels of detail that may require shifting throughout programming. For example, during a course in Ancient History, one may correctly historically sequence Herodotus' History before Thucydides' Pelopennesian Wars. Within the Thucydides' sequence, however, one may have to use the psychological principle of discrimination to aid the student in distinguishing between "expediency" and "principles" emphasized by Thucydides. In other words, the linking together of content in a learnable sequence may require the interweaving of several principles at more than one level.

Empirically, matters may be different. After the course has been tested, it may have to be re-sequenced based on the data received. Discussion of such a confrontation will be examined later in this paper (see Paragraph 2.4.2).

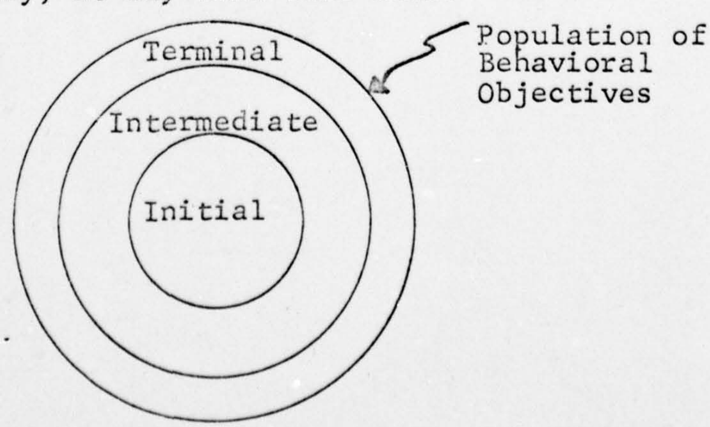
2.2.2 Ordering of Initial and Intermediate Behavioral Objectives (6.0)

The content sequences are ordered isomorphically with the general objectives of the course as has been discussed in Paragraph 2.2.1. In a similar way, initial and intermediate objectives are ordered isomorphically within a particular content sequence. For example, if mechanical linkages must be taught before electronic linkages according to the content sequence principle selected (in this case, logical) then initial and intermediate behavioral objectives relevant to mechanical concepts must be ordered WITHIN the mechanical content sequence. The initial and intermediate behavioral objectives relevant to electronic concepts therefore, must be ordered WITHIN the electronic content sequence. In another example from the teaching of heredity: The initial and intermediate behavioral objectives regarding the concept of non-

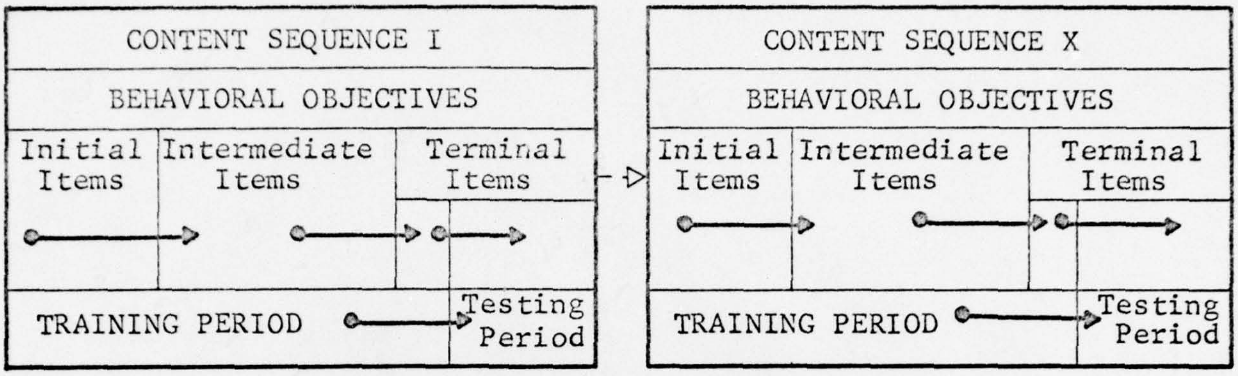
linkage (independence) of assortment of genetic material, according to Mendel, might be content sequenced before the intiial and intermediate behavioral objectives regarding the concept of linkage (non-independence).

What, in the latter example, would be the "initial behavioral objectives?" Certain concepts from logic and descriptive statistics for example, are important in teaching hereditary theory. These would be the initial behaviors. The purpose of differentiating and ordering the initial and intermediate objectives is to prepare for the next major phase, Phase III, Preparation of the Programmed Course. Hence, with the initial behavioral objectives selected, the programmer can proceed without losing track of what he is doing. The initial behavioral objectives would be relevant statistical concepts (not the whole of descriptive statistics), and relevant logical concepts, (not the whole of elementary logic). Initial behavioral objectives link the "easiest" part of a given subject matter with the prerequisites for a course. Experimentation with the course will determine whether the initial objectives were the easiest for the students.

Intermediate behavioral objectives within a content sequence would be the application of the relevant logic and statistics to genetic problems. It is important to realize that initial and intermediate behavioral objectives are part of the population of behavioral objectives which includes the terminal objectives; however, the sampling process is not random. It is purposeful and isomorphic with the content sequence being established. Logically, it may look like this:



Sequentially, the situation may look like this:



The filled arrows indicate the order of sampling within a particular content sequence. The unfilled arrow between sequences indicates omitted content sequences. Note that some of the terminal items may be given during training and some during testing. It is important to note that the terminal behavioral objectives given during training are identical in kind to the terminal behavioral objectives given during training. Thus, if a training terminal objective selected was:

$$y = \tfrac{1}{2} x + b \qquad (0 \leq x \leq 10;\ b = 5)$$

$$y = ?$$

then a testing terminal objective could be:

$$y = \frac{x}{2} + b \qquad (0 \leq x \leq 100;\ b = 10)$$

or, if the training terminal objective selected was:

all A's are B's
all B's are C's
―――――――――
all A's are ?

There a testing terminal objective selected could be:

All Cats are Felines.

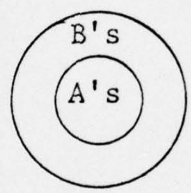
All Felines are Animals.

All Cats are ?

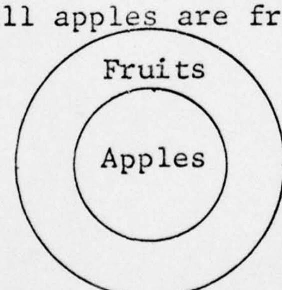
This situation is not true, however, for initial and intermediate behaviors. The essential difference between these two objectives is that the degree of difficulty is higher for the intermediate behavioral objectives. To continue the sample from logic illustrated above, one might write an initial behavioral item as:

"All A's are B's,"

and illustrate it with a Venn or Euler diagram as follows:

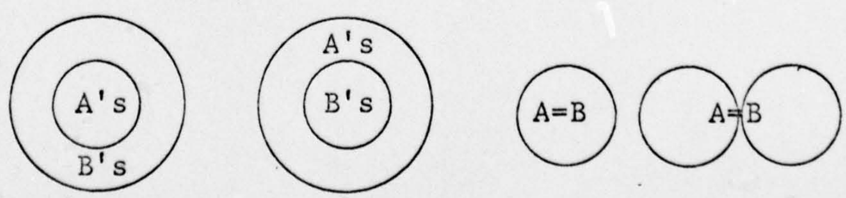


or "All apples are fruits."



The intermediate behavioral objective would be written:

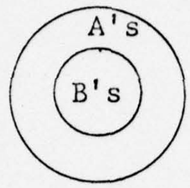
"All A's are B's." Which diagram illustrates this?



*a. b. c. d.

*correct answer

or: "Which of the following statements is correct from the diagram?"



a. All A's are B's
b. No B's are A's
*c. All B's are A's
d. All A's equal B's.

Clearly this arrangement seems intuitively more difficult than merely pointing out to the student a verbal statement and its illustration by a diagram. Hence, the verbal statement and its illustration would be an example of an <u>initial</u> behavioral objective while the task of translating an illustration into a verbal statement would be an example of an <u>intermediate</u> behavioral objective. In the final analysis; however, it is necessary to experiment with the objectives to ascertain whether they are initial or intermediate.

2.2.2.1 <u>Ordering Behavioral Objectives</u>: <u>Intuitive and Empirical</u>.

It is necessary to state that the order of difficulty of behavioral objectives intuitively proposed by the subject matter expert and/or programmer may turn out to be different from that encountered by the students in training. The number of students making the various choices on each alternative to a particular behavioral objective is an empirical index of item difficulty. This index can be expressed in probability terms when the total number of students selecting a particular choice is divided by the total number of students making all responses to all alternatives for a single item. ("Item" means behavioral objective, initial, intermediate, or terminal, plus a question.)

*correct answer

This measure is not enough, however, to determine the evaluation of the total frame which includes new information as well as the behavioral objective to be taught. In order to evaluate a total frame, the subjective confidence that the learner has in his answer is required. For each item, as illustrated in Section 1.0 of this paper, there would be a distribution of subjective probabilities for each student for every item: the confidence he has that he has selected the right alternative based on what he has learned up to and including the particular frame of interest. For example:

Frame 101. What is the value of $y = e^x$ when $x = 0$?

a. .5
*b. 1.0
c. 0
d. e

Now, the empirical probability of error might be distributed as follows: (N=30)

a. 1/15
*b. 3/5
c. 1/6
d. 1/6.

However, the subjective probability distributions based on the confidence of their answers by students might be:

	SUBJECTIVE PROBABILITIES					
ALTERNATIVES	.00	.25	.50	.75	1.00	Σ
a.	0	1	1	0	0	2
*b.	0	6	6	5	1	18
c.	1	0	1	3	0	5
d.	1	0	3	1	0	5
Σ	2	7	11	9	1	30

*correct answer

Now the subjective probabilities distributions show that some students will be surprised that their "guesses" or "lack of learning" were wrong. Forgetting or inattention are also possibilities as causes of errors.

More important, however, this distribution, along with the error distribution, might suggest to the experimenter-programmer that either the frame is not written properly or that this frame is written properly but that it is too difficult; hence, the content sequence is wrong when adjacent frame probabilities are examined in conjunction with it. There even may be a "run" of "bad frames." However, for any one student's record, analysis may show "bad habits": carelessness, guessing, etc.

The point of adjustment of a frame depends on which alternative had the greatest number of subjective probability choices and was also the wrong answer. The reader should be reminded that the two probability indices (objective and subjective) would NOT be used or required of a student after the programmed course had been trial tested. These indices are used as analytic devices only during experimentation. These indices are suggestive only and a more rigorous mathematical and logical treatment will be presented in a forthcoming paper.

A further reminder is necessary. The subjective probabilities predicted by the student about the correctness of his answer during pre-testing of the course and any reinforcement system used during learning or actual use of the course may be independent of each other. The probability of reinforcement may be 1/4, 1/5, 1/2, 1/3, etc., depending on the number of alternatives available but the number of points gained or lost depends on the selection of the correct answer or the wrong answer. For example, for a four choice item, if "b" is the correct answer, then "b" may equal +10 points. However, wrong answers may subtract 2, 3, or 5 points depending on the seriousness of the error. Careless errors may subtract 2 points. Forgetting of a previously taught frame may subtract 3 points while forgetting a prerequisite or similar error may subtract 5 points. (This is a sample reinforcement system).

The statistical reliability of items and tests is a well-known discipline and appropriate texts* should be consulted to see that relevant statistics are used to prepare items and tests in final form.

2.3 Phase III: Preparation of Programmed Course

During this phase, the first draft of the programmed course is written and practice items (behavioral objectives) are expanded. The programming must be aligned with the taxonomy of general objectives, the content sequence, and with the facts of learning.

2.3.1 Identification of Psychological Skills (8.0)

Each behavioral objective must be analyzed in order to determine the psychological skill or skills involved. For example, if discrimination between a fractional and an integer exponent must be made, then discrimination is the chief psychological skill to be taught regardless of the content sequence or taxonomy involved. Certain facts about discrimination** are known in training for this skill and these should be reviewed prior to actual programming. The Taxonomy of Bloom will provide the programmer with a technique for writing the terminal behavioral objectives, but the psychological skill to be taught focuses on the information given the student before he is given a practice item.

For example, to give the student information that the "x" in $y = e^x$ is an exponent and then present him the following practice item:

$$y = e^{\frac{1}{2}} = ?$$

*Guilford, J. P., Fundamental Statistics in Psychology and Education. New York: McGraw-Hill, 1950.
**Blaisdell, F. J., op. cit., Section 1.0.

As a student practices the complete course he should gain a higher score each time until he can complete at least two successive practice sessions with at least 90% or higher score based on the total number of points possible.

2.2.3 Specification of Reliability for Behavioral Objectives (7.0)

From an analytic viewpoint, as the student practices he should make fewer errors each practice session so that the probability of his responding correctly should approach 1.00 on each frame. If a constraint is placed on the programmed course in such a way that it must be re-written if 90% of the students do not achieve 90% or higher on the terminal objectives (on a final examination) then it follows that some index of reliability for the individual frames must be imposed while the student practices.

The reason that individual learner reliability is important is that 90% may be achieved on the whole course after practice, but some frames may have been learned at a lower level. It is also important to remember that this reliability is not the statistical reliability of items, but the reliability of the learner. Hence, each student's record should be examined to see whether with each practice session more and more individual frames are being answered correctly.

To do this a programmed course could not be divided into initial, intermediate, and terminal sections because a terminal behavioral objective may occur early or late in a course. It appears better to specify for all behavioral objectives the same level of reliability. An example would be to specify two successive practice sessions with 100% score on all frames. Three or more successive sessions with 100% scores could be specified depending on the length of the period between end of practice and testing or upon the error tolerance allowable on the job.

It is unfair if he doesn't know that "x" could be an integer, a fraction, a decimal fraction, a mixed decimal, or mixed fraction. Such discriminations must be carefully made known in a conscious way to the student.

Each frame should be identified in every way. For example, Frame 2.3.1: "Discrimination, exponents," in addition to its taxonomic and content sequence identifications. This is merely a matter of bookkeeping after decisions about each frame has been made.

2.3.2 Programming of Psychological Skills (9.0)

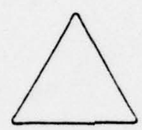
Programming itself is probably the most difficult and time-consuming step in programmed instruction. The aim in this paper is to systematize the instructional process. The programming process is the most elusive process to systematize. Previous steps have lent themselves to systematization in such a way as to place a constraint on careless programming. The specification of behavioral objectives in terms of Bloom's Taxonomy, the sequencing of content based on identifiable principles, the ordering of the behavioral objectives isomorphic with the content sequence, the identification of psychological skills, and the specification of reliability criteria tend to prevent the programmer and/or subject matter expert from introducing material out of sequence or irrelevant behavioral objectives.

It will be very unlikely that the programmer can sequence and program material correctly prior to experimentation with the programmed course. Needs of students differ so that the branching process will finally depend upon the kind of errors students make most frequently. It is necessary to distinguish again the initial tryout of a draft of the programmed course and the experimentation or testing of it in a formal experimental design.

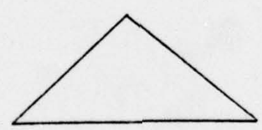
2.3.2.1 General Programming Format

The programmer should cross-index all frames and all alternatives so that he doesn't get lost in a program. When a programmed course is stored on the computer this matter of bookkeeping must be maintained.

Besides the bookkeeping aspect, the most important step is to present information to the student before introducing any test material (behavioral objectives). The key to the information presented is the psychological skill to be taught as was stated in step 2.3.1. For example, if an equilateral triangle,



is to be differentiated from an isosceles triangle,



it is necessary to present in detail how they are different and how they are similar. Discrimination is the psychological skill to be taught.

It is very unwise to present a question which implies an implicit assumption or rule unless the assumption or rule has had a previous explicit presentation. Subtle differences in forms or formulas may require very careful explicit presentations in several ways and be followed by several examples. In special cases where word endings and their associated grammatical rules are to be memorized, as in foreign languages, examples must be provided for all cases and their combinations at least once.

If a generalization is the psychological skill to be taught, the principle and examples of it must be presented in different ways so that it can be identified reliably during training and rapidly during testing. For example, it is not enough to present the formula,

$$e^{o} = 1$$

and expect a student to generalize to:

$$y^{o} = 1, \ 2^{o} = 1, \text{ etc.}$$

If problem solving is a skill to be taught, then factors which bear on the particular problem at issue must be explicitly and thoroughly illustrated. Such statements as: "The proof will be left as an exercise" and "It is easy to prove that...." must be avoided. The proof should be given in detail. If a proof is easy or difficult, both cases should be presented explicitly where needed. Each detail rule or "lemma" should be stated and shown where it applies. If a formula requires similar units to be used before it can be applied, then all possible examples should be given--unless some examples will not be included in the population of behavioral objectives.

Do not expect students to analyze and remember a long SORITES ~~series~~ consisting of truncated syllogisms. Such situations often occur in the analysis of poetry and plays. Backgrounds of students vary so much that drawing room language for some students will be perfectly clear while barroom language may be utter nonsense to those same students. Meanings of these words should be available to the learner to prevent guessing and misinterpretation.

Wherever a number of psychological skills are combined in a behavioral objective as for Comprehension*, Analysis*, Synthesis*, each skill involved should be taught explicitly before

*see Bloom, op. cit.

it is combined into the objective. "Good Gestalt," however, should be maintained. For example, if a poem or musical piece is presented, the whole of it should be given to the student. However, analysis of the work should be given in such a way that its parts are made explicit as well as its whole structure.

If the skills involve man-machine interaction, then a task analysis should be made first so that all psychological skills are known to the programmer and the time requirements, both simultaneous and sequential, are available. A task analysis should relate all interacting subsystems, including the personnel subsystem, as these subsystems interact through time to produce mission phases in which personnel must perform tasks. Other constraints besides time should be known to the programmer. All relevant inputs and outputs to and from the learner must also be known to the programmer. The expected reliability for each task must be stated: e.g., " $\lambda = \frac{2}{100}$ " or "No more failures (λ) than two in 100 attempts allowable for Task 12.2.1 to be performed twenty times per hour." Such requirements identify explicitly for the programmer the criteria that must be achieved by a student trained by the programmed course.

2.3.3 Expansion of Practice Frames (10.0)

Once a frame has been written for a particular objective, it should be used as a model for subsequent frames which will be used as practice problems. These frames will be brief and will indicate why and how a particular alternative was correct. This information will be followed by a similar problem which introduces no new teaching points or psychological skills. The number of frames devoted to practice problems is a function of the weighted importance of the behavioral objective. For example, in a course in a foreign language where poetry analysis is the general objective, more practice would be given perhaps on adjectives and the subjunctive mood than on prepositions and the indicative mood.

When computer-assisted instruction is integrated with simulators or trainers, all the steps outlined in this paper would be required. Caution should be exercised in directly applying the steps primarily designed for symbolic behaviors to perceptual-motor skills. Careful check of the Handbook of Human Learning* or similar documents** should be made while programming to determine whether additional information should be inserted or whether certain steps should be altered. For example, when a console is a part of a larger system and when the operator must combine display information and control status to alter the control positions to obtain a changed display, the operator can learn the system operations to perform the mission by using a programmed simulator integrated with CAI. Here the learner presses buttons and moves other controls as well as makes discriminations and generalizations from the displays. The timing factor, required by the mission, may add a constraint not found, say, when simply studying algebra for a future examination. For example, an operator's tracking varying symbols on radar or sonar during a HUK-type mission may or may not have the threat vector during training as in an actual situation; however, the level of skill to be achieved and the proper organization of responses require observance of all the steps discussed so far.

2.4 Phase IV: Pre-test of Programmed Course

The activities carried out during this phase involve pre-testing the programmed course and revising it based on data from a very small sample of persons. It is inadvisable to computerize a programmed course prior to de-bugging it.

2.4.1 Pre-testing the Initial Programmed Course (11.0)

The first step in pre-testing the initial draft of the programmed course is to administer the selected prerequisite test to a group of students. The prerequisite test consists of terminal

*See Blaisdell, F. J., op. cit.
**Morgan, C. T., Chapanis, A., Cook, J. S., and Lund, M. W., Human Engineering Guide to Equipment Design. New York: McGraw-Hill, 1963.

behavioral objectives from a previous course. The prerequisite test does NOT contain items to be taught in the programmed course; however, the items contained in the test should be those necessary and relevant to learning the programmed course. The prerequisite test contains a representative sample of the terminal behaviors from a previous course of study and should have been standardized by recognized psychometric methods*.

The prerequisite test scores for the pre-test group of students should be reviewed and those students who do not score 90% or higher should not be allowed to take the tentative programmed course.

The second step in pre-testing the programmed course is to administer the preliminary draft of the programmed course to the students making 90% or higher on the prerequisite test. It is wise to make careful notes of all comments students may make in regard to being trained by the programmed course. The students should be encouraged to write on the margin of the frames any inaccuracies or difficulties they may meet.

Students should be required to indicate in the programmed text the answer they chose in the following order: 1^{st}, 2^{nd}, 3^{rd},...,n^{th}. The student should be informed that the course is being pre-tested and that this information is important.

Students should be paid for their efforts. A scale of remuneration should be proposed to enable the fastest student who completes the programmed course and makes 100% correct on the terminal objectives will receive the most dollars. Ties will be similarly rewarded.

The third step in pre-testing the draft of the programmed course is to administer after training an examination consisting of a representative sample of terminal objectives. This examination

---

*Adkins, D. C., et. al. Construction and Analysis of Achievement Tests. U. S. Government Printing Office, 1947.

is not to be standardized in the usual sense*. The terminal objectives ARE the behavioral responses desired; hence, statistical averages representing a hypothetical individual from whom all others are scored is not relevant.

2.4.2 Revision of Pre-test Materials (12.0)

Based on the time to train and to test, errors made on the course and test, and comments received from the students, the programmed course should be revised frame by frame and the test item by item.

Since the materials prepared up to this step are preliminary in nature, the analysis of the results should rely less heavily on statistical or other psychometric techniques and more heavily on comments and error accumulations on the examination items and programmed frames. The scheme of using subjective probabilities might be tried as suggested earlier. (A future paper will treat this approach in a more rigorous fashion.)

2.4.2.1 Preparation of Branched Frames

Based on errors and comments of students, additional branched frames should be prepared or extended especially at points where errors have accumulated on test or course and it is considered inadvisable from other evidence (IQ, aptitudes, etc.) to change a particular frame but to use more branching frames.

2.5 Phase V: Preparation of Computer Program and Course Storage

During this phase the computer program is designed in accordance with the requirements of the programmed course, the prerequisite test and the terminal examination. Appropriate input-output devices must be selected to meet course requirements

*Adkins, D. C., op. cit.

and computer storage limitations. During this phase the programmed course and examinations are stored. A brief pre-test of the stored materials are run and necessary revisions made in the computer program prior to beginning experimentation proper.

2.5.1 Design Computer Program for the Course (13.0)

The design of the computer program is a function of the design of the programmed course. The course requirements act as constraints upon the computer program. A clear distinction must be made between what must be taught and what can be designed, hardwarewise, to help teach and control behavior.

It should be kept in mind that the system methodology proposed in this paper is aimed at producing a standard instrument with known parameters to investigate human learning and retention. The required display-control devices and a computer are used to assist in the investigation.

Since computer programming design is a well known art and since this paper is concerned primarily with the psychological aspects of human learning and retention, details of computer program design can be found in other publications covering such design.

2.5.2 Computer Program Design for Tests (14.0)

Mental test techniques have been in existence for several decades. The design of a computer program which presents the items one at a time and records the student's responses is already within the state-of-the-art. Feedback to the examinee as to how he is doing on an examination is not customary. Feedback data required for the examiner and experimenter is discussed in Section 2.6.2, Data Recording.

2.5.3 Input-Output Devices Selection (15.0)

This paper is not concerned with display and control devices as such but only with their general specification for a particular research program in computer-assisted instruction. In

general, the learner may receive an optical or a printed output one frame at a time. He should be able to obtain a previous frame by an appropriate control. He should be able to have displayed to him the results of his decision, his accumulated score, the time he is consuming per practice session, the total time to learn to criterion, and the time between onset of a frame and his response. Dynamic displays and controls are not required for the experiments anticipated at this time. Further information outputs pertaining to the experimenter are discussed in Section 2.6.2, Data Recording.

2.5.4 Storage of Course and Examinations (16.0)

This is a step in the computer phase of the methodology. During this phase, the experimenter monitors the process and furnishes information where required.

2.5.5 Testing of Stored Course and Examinations (17.0)

As for Step 2.5.4, the testing of the stored course and examinations is a computer phase and does not indicate testing of learners. Details pertaining to this phase are covered in well-known sources of computer technology.

2.6 Phase VI: Administration of Pre-test and Course

During this phase, selected students will take the pre-test and then begin their instruction with the computerized course. Data concerning inputs to the students and their responses will be recorded on a realtime basis.

2.6.1 Administration of Materials to Appropriate IQ Levels (18.0)

Students should be selected for training in such a way that the sample is representative of the population that is likely to be taught by such a course. If the programmed course is 5th grade arithmetic, then a representative sample of 5th graders should be selected. Variables other than IQ may be included in the sample such as class and "race," native or foreign birth, etc.

The prerequisite test is administered prior to the programmed course and the scores are tabulated and analyzed. If a student does not achieve 90% on the pre-test then he should be excluded from the programmed instruction.

2.6.2 Data Recording (19.0)

Figure 2-1, Functional Flow Diagram, lists all data that should be collected from the prerequisite test, the programmed course, and final examination. For convenience to the reader, the following is a repetition of the list:

- Student responses by alternative
- Time between responses
- Number of errors
- Number of correct responses
- Cumulative errors per practice session
- Errors per frame
- Total time to reach practice criterion
- Total errors before and after reaching criterion
- Number of errors per frame for all students
- Confidence levels distribution per alternative per student (these are the subjective probability estimates).

It is necessary to record the following data by computer as each student takes the prerequisite test:

- Item number
- Response alternative
- Response correct
- Total items correct
- Total items wrong
- Total test time

It is not always necessary to give the prerequisites test by computer; however, for later calculations it is easier to record and store this data by computer as well as to administer the prerequisites test by computer.

2.7 Phase VII: Administer Objectives

During this phase, the terminal objectives will be administered and responses will be recorded and analyzed.

2.7.1 Administration of Terminal Objectives (20.0)

The procedure for administering the terminal objectives (final examination) by computer is similar to that for administering the prerequisites test. Both tests can be administered manually. Real-time data should be recorded as for the prerequisites test when administered by computer. The terminal objectives should be administered immediately after original learning. (see Section 2.8.1).

2.7.2 Score and Item Analysis (21.0)

The technology for making the analyses of scores and item analysis can be found in well-known texts on these subjects*. It is easier, however, to have the computer make the recordings, as described in Section 2.7.1, as an aid in the analysis of results.

2.8 Phase VIII: Re-administration of Terminal Objectives

During this phase, the terminal objectives only will be administered 30 days and 60 days after the initial administration of these objectives and appropriate comparisons made with original specifications made during Phase II, Functional Step 7.0.

2.8.1 Re-administration of Test After 30 Days (22.0)

The same subjects will take an examination composed of items selected from the population of terminal objectives. To control for the memorization of test items, parallel tests will be constructed and administered. Data will be recorded as for the test given immediately after original learning.

*see Adkins, D. C., op. cit., and Guilford, J. P., op. cit.

"Immediately after original learning" will be defined to mean no sooner than 24 hours after original learning and no later than 48 hours after original learning. The time elapsing between original learning and examination will be recorded.

Control for interpolated activity is not contemplated by the procedures described in this paper, although modification of the methodology to permit examination of this important factor as well as fatigue, homogeneity of material, etc., could be made. Future papers will develop procedures for such experiments using the present methodology.

2.8.2 Re-administration of Test After 60 Days (23.0)

The procedures described in Section 2.8.1 will be used to administer the third examination after original learning. This examination will be parallel to the others to test in order to prevent item memorization. Appropriate data will be collected as for other examinations.

2.8.3 Reliability Comparisons (24.0)

After each test administration, the required reliability achieved during training by the student will be compared with the reliability on each examination. The following might be an example of the students' behavior:

END OF TRAINING:	100% Correct, 5 Trials Required
1st EXAMINATION:	95% Correct
2nd EXAMINATION:	92% Correct
3rd EXAMINATION:	90% Correct

In this case, both the programmed course and examinations appear to be well prepared. Total trials to learn appear not to be excessive in meeting the criterion of two successive practice sessions with a score of 100% correct.

However, if the students' behavior showed the following:

END OF TRAINING:	100% Correct; 25 Trials Required
1st EXAMINATION:	80% Correct
2nd EXAMINATION:	70% Correct
3rd EXAMINATION:	65% Correct

Then it appears that considerable revision of the entire course is required. The revision should not always begin during this phase. It would have been better to have noticed if excessive trials were needed to achieve the training criterion before giving the examination. (The initial draft of the course may not have revealed the same results as the experiment.)

Speculations concerning the outcomes of this methodology are not in order in this paper. It was the purpose of the paper to outline an integrated methodology to attack systematically the functional steps in exploring human learning. Considerable revision and extension of the methodology is quite likely, especially with regard to providing operational ways of meeting student-centered instruction. Early discussion of student-centered instruction has been proposed by Rogers[*] and an early attack on the problem has been suggested by Mager[**]. A suggested method using subjective probability has been proposed in this paper but not in a rigorous manner; however, a future paper will present a more logical and mathematical treatment of student-centered content sequencing and programming.

*Rogers, Carl R., Client-Centered Therapy. Chapter 9, Student-centered Teaching, New York: Houghton Mifflin, 1951.
**Mager, R. F., On the Sequencing of Instructional Content. Psychological Reports, Southern Universities Press, 1961, Palo Alto, Calif.